

센서 네트워크에서 집계 연산을 위한 점진적 뷰 관리¹⁾

최주리* 이민수

이화여자대학교 과학기술대학원 컴퓨터학과

blueri00@ewhain.net* mlee@ewha.ac.kr

Incremental view maintenance for aggregation operator in sensor networks

Choi Juree Lee Minsoo

Dep. Computer Science and Engineering, Ewha Institute of Science and Technology,
Ewha Womans University

요 약

센서 네트워크에서 집계(aggregation) 연산은 센서 네트워크를 데이터베이스로 구현하는데 있어서 중요하게 제공되어야 하는 서비스이다. 현재 연구되고 있는 것으로 집계 연산을 센서 네트워크의 특징상으로 분류하여 근접한 결과값을 받는 것을 허용하고 집계 값을 자식노드가 부모노드로 보내는 기간을 부모노드가 자식노드에게 나누어 할당하여 센서 네트워크상에 적절히 구현하는 것에 대해 알아보고 집계 그루핑을 하는 과정에서 데이터웨어하우징 연구의 최신기술인 점진적인 뷰 관리 기법을 통해 센서노드들의 평균값에 대해 새로운 값이 추가될 때 다시 모든 값을 계산하기 않고 변경된 값만 적용하여 계산함으로써 좀 더 에너지 효율적으로 확장하는 것을 제안하였다.

1. 서 론

최근에 센서 네트워크를 데이터베이스로써 구현하려는 여러 시도들이 있어왔다. 대표적으로 코넬대학에서 수행하고 있는 COUGAR 프로젝트[1]와 UC Berkeley에서 연구하는 TinyDB[2]등이 있다. 여기서는 센서 네트워크에서 크게 고려해야 하는 에너지 효율성을 위해 중앙집중적인 데이터베이스가 아닌 네트워크상에서 질의를 던지는 구조로써 설계를 하고 있다. 즉, 센서 노드에서 모니터링 하는 데이터들에 대해 네트워크 상으로 질의를 주어 결과를 받는다.

이때 실제로 센서 네트워크에서 알고자 하는 값들이 센서가 읽어들이는 순수한 데이터 보다는 데이터의 평균, 최대값 등 집계(aggregation)값을 필요로 하는 경우가 많다. 그래서 이 논문에서는 기존의 데이터베이스 연산자 중 집계값에 주안점을 두고 집계 그루핑(aggregation grouping)에 대해 데이터웨어하우징에서 연구되고 있는 점진적 뷰 관리(incremental view maintenance)를 적용시켜 센서노드의 메모리 한계를 줄이고 에너지 효율적인 방법을 제안할 것이다.

2. 관련 연구

1) 이 논문은 2004년도 두뇌한국21사업에 의하여 지원되었음

TinyDB를 연구하고 있는 UC Berkeley에서는 계산능력을 가진 motes라는 조그만 센서 장비를 개발하였으며 TinyOS라고 불리는 운영체제 시스템을 사용하여 ad-hoc 네트워크상에서 motes의 개발이 용이하게 되도록 디자인되어진 기본 기능들을 제공하고 있다. 그리고 에너지 효율성을 위한 집계 연산 서비스인 Tiny AGgregation(TAG)를 제안하고 있다.

TAG에서는 에너지 효율성을 높이기 위해 집계 연산들에 대해 근접한 집계 결과값을 제안했는데, 이는 온라인 모니터링에 유용하고 통신 비용을 줄일 수 있으며 네트워크상의 에러 복구에 대한 것에 미연에 방지할 수 있다.[3] TAG의 핵심 알고리즘으로는 네트워크 상으로 집계 질의(query)가 센서 노드들에게 던져진 후, 자식 노드로부터 부모노드로 집계값이 계속적으로 전송이 되는데, 이러한 작업은 결과값을 내야하는 epoch이라는 기간동안 이뤄지고 이는 마지막 노드에서 결과값을 받지 못했을 경우 많은 메시지가 소요되게 된다. TAG에서는 이를 부모노드가 자식노드에게 epoch을 나누어서 할당하여 수행함으로써 통신비용을 줄일 수 있다. 그렇다면 부모노드가 할당하는 기간을 선택하는 기준이 필요한데 이는 정확한 값을 받을 때까지 기다리게 되면 많은 주파수가 필요로 하여 전력을 많이 소모하게 되고 또한 네트워크 토폴로지에 대한 전략에 따라 달라지기 때문에 좀더 연구가 필요할 것이다. 그리고 TAG에서는 SQL문의 GROUP BY와 같은 그루핑을 제공한다. 질의를 자식노드에게 던질 때에 그룹을 분류하는 술어(predicate)을 같이 전송하여 자식노드가 그룹을 선택하고

해당 그룹 번호를 태그로 붙여 부모노드에게 돌려주게 된다. 또한 HAVING절 역시 그루핑 질의와 같이 보내줘서 미리 필터링된 결과를 돌려줄 수 있다. 이외 TAG에서는 성능과 정확도를 향상시키기 위한 여러 최적화 기술들을 제안하고 있다.

3. 본 론

3.1 점진적 뷰 관리

데이터웨어하우징 분야의 데이터베이스 성능 향상을 위한 대표적인 기술인 실체화 뷰(materialized views)는 일종의 인덱스와 유사한 역할을 수행한다. 일반 뷰는 다른 릴레이션으로부터 유도된 릴레이션(derived relation)으로써 사용자 하여금 복잡한 질의를 간단하게 표현하는 역할을 하고 실제 질의 처리는 해당 뷰의 기본 테이블에 대한 질의로 재작성된 후 수행된다. 반면, 실체화 뷰는 해당 뷰의 정의에 해당하는 실제 내용이 디스크에 저장되어서 기본 테이블을 대상으로 한 질의에 대해 가능한 경우에 빠른 질의 처리를 위해 실체화 뷰로 재작성해서 질의 처리를 수행한다.

이러한 실체화 뷰를 관리하는 방법으로 점진적 뷰 관리 기법은 기본 릴레이션의 변경 사항만을 이용하여 뷰를 관리하는 방법이다[4]. 이 방법은 소스 데이터의 갱신(삽입, 삭제, 수정)에 대해 데이터 소스의 접근 없이도 뷰에 변경 사항을 반영하는 뷰의 자체적 관리 기법이다. 이러한 뷰의 자체적인 관리는 단지 실체뷰의 내용과 기본 릴레이션의 변경 사항만을 가지고 뷰를 관리한다[5][6]. 따라서 갱신된 데이터 소스에 대한 직접적 접근이 없기 때문에 질의에 대한 빠른 응답과 높은 이용 가능성을 제공하므로 대량의 뷰들을 관리하는데 매우 효율적이다. 대부분의 뷰 관리 기법은 이러한 뷰에 대한 변화를 계산하기 위해 수학적 식으로써 뷰를 정의하고 변화를 표현한다. 표1은 점진적 뷰 관리 기법을 표현한 것으로 기본 릴레이션의 변화값인 t 만을 구해서 뷰에 적용하는 식이다.[6] ∇R 은 R 테이블에서 삭제가 된 변화값이고 ΔR 은 삽입된 변화값이다.

표 1

$$t = \{R_1 \leftarrow (R_1 \div \nabla R_1) \cup \Delta R_1, \dots, R_n \leftarrow (R_n \div \nabla R_n) \cup \Delta R_n\}.$$

집계 함수를 예로 들어보자. S라는 관계 테이블이 있고 한 건물안 여러 방의 온도값을 가지고 있다. 다음은 여러 방의 온도값 평균값을 정의한 뷰인 V이다.

$$V := AVG(S) = TOTAL(S)/COUNT(S) = \sum_{id}(S) / \sum_1(S)$$

집계 연산을 이용해 평균값(AVG)은 SUM과 COUNT를 이용해 구할 수 있으며 COUNT는 1일 때는 항상 1을 돌려주므로 \sum_1 로 표현하고 SUM은 \sum_{id} 로써 표현하자.

여기서 S 테이블에서 변화값을 구하기 위한 최소한의 계산은 $S_n = (S \div \nabla S) \cup \Delta S$ 가 되고 이를 뷰에 적용하기 위해서는 다음과 같이 계산할 수 있다.

$$AVG(S_n) = (\sum_{id}(S) - \sum_{id}(\nabla S) + \sum_{id}(\Delta S)) / (\sum_1(S) - \sum_1(\nabla S) + \sum_1(\Delta S))$$

이는 뷰를 기본 릴레이션의 변화값이 생길 때 다시 전체적으로 계산할 필요없이 변화된 값만 추가하여 변경할 수 있게 한다[5][7].

3.2 그루핑

센서 네트워크의 특정한 지역에 있는 많은 노드들에게 그림1에 나와있는 질의를 던진다고 하자. 여러 노드들로부터 빛의 세기의 평균값과 온도를 받고 온도값으로 그루핑을 하기 원할 때 질의와 같이 그룹을 결정하는 술어(predicate)가 같이 보내진다. 질의를 받은 자식노드는 자

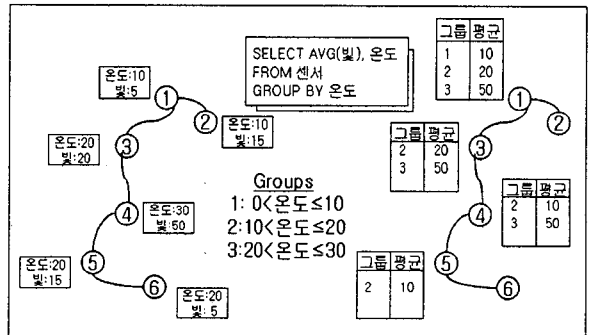


그림 1

신이 속한 그룹을 선택하여 그룹 번호를 태그로 붙여 부모 노드에게 보내고 되고 부모노드에서는 자신이 가지고 있는 그룹과 비교해 저장되어 있는 그룹일 경우 평균값을 내고 새로운 그룹일 경우 추가한다.

그런데 센서 노드의 경우 메모리가 작다는 한계성을 가지고 있기 때문에 많은 그룹에 대한 값을 가지고 있기 힘들고 이러한 경우 기존 데이터베이스 알고리즘인 partial preaggregation을 적용해 어떤 그룹을 희생시키는 방법으로 오버헤드를 줄일 수 있다[8].

그리고 같은 그룹내에서 평균값을 구하는 경우 한 노드의 값이 변했을 경우 부모노드는 자식 노드에서 그룹 번호와 값을 받아서 다시 평균값을 계산하게 된다. 이러한 경우 모든 값을 받아서 평균값을 다시 계산할 경우 센서 노드의 처리하는 시간이 많이 걸리게 되어 수명이 줄어들게 된다. 그래서 점진적 뷰 관리 기법을 이용해 변화된 값만을 반영해 평균값을 계산하여 센서 노드의 메모리, 시간, 전원 등을 효율적으로 이용할 수 있는 방법이 된다. 다음과 같이 질의에 대해 r이라는 뷰를 정의하고 같은 레벨에 있는 여러 센서로부터 값을 받아 한 노드에서 평균값을 구하는 경우

를 생각해 보자.

```
create view r(빛, 온도) as
(select avg(빛), 온도
 from 센서
 group by 온도)
```

여기서 r이라는 뷰는 센서1,센서2,..센서n으로부터 온도값과 빛의 평균값을 계산하고 온도값으로 그루핑이 되어져 있다. 여기서 n개의 센서중 센서2의 빛의 세기가 변했을 때 기본 릴레이션(base relation)인 센서2의 변화값만을 가지고 평균이라는 뷰에 변환된 $\Delta(\text{평균})$ 값을 계산해 준다.

$$\Delta(\text{평균}) = \text{센서1} \cup (\text{센서2} - \nabla \text{센서2} + \Delta \text{센서2}) \cup \text{센서3} \dots \cup \text{센서n}$$

즉, 다시 계산할 필요없이 점진적 뷰 관리 기법을 이용하여 평균 $\cup \Delta(\text{평균})$ 을 적용하여 센서 노드의 빛의 평균값을 에너지 효율적으로 계산해 볼 수 있다.

4. 결 론

이 논문에서는 센서 네트워크를 데이터베이스로 구체화하는데 있어서 질의를 네트워크상으로 던지고 결과값을 받는 구조를 기반으로 하였다. 그리고 실제로 구체화 하는데 있어서 기존의 데이터베이스 연산자중 집계(aggregation) 연산이 센서 네트워크에서 핵심 서비스로 제공될 필요성이 있으며 센서 네트워크에서 에너지 효율적으로 구현을 하기 위한 방법을 기존의 연구되었던 TAG를 기반으로 하여 연구해 보았다.

특히 데이터웨어하우징 연구에서 중요한 기술인 뷰 실체화, 그리고 점진적인 뷰 관리 기법에 대해 알아보았고 이를 센서 네트워크의 집계 함수 그루핑에 적용을 해보았다. 이는 센서 노드의 메모리 한계를 줄이고 에너지 효율성을 높일 수 있는 방법이다.

앞으로 그루핑에서 HAVING절을 구체화 하는데 있어 좀더 에너지 효율적인 방법을 연구해 볼 필요성이 있으며 이 논문에서 제안한 점진적인 뷰 실체화를 그루핑에 적용 시킨후의 성능향상에 대한 연구와 여러 최적화 기술들에 대한 연구가 좀더 필요하다.

5. 참고 문헌

- [1] Y.Yao and J.Gehrke. "The Cougar approach to in-network query processing in sensor networks" SIGMOD Record, 31(3):9-18, 2002.
- [2] S.Madden et al. "The design of an acquisitional query processing for sensor networks" In Proc. of SIGMOD, p491-502, 2003.
- [3] Samuel Madden, Michael J. Franklin, and Joseph M. Hellerstein Wei Hong. "TAG:a Tiny AGregation

Service for Ad-Hoc Sensor Networks", ACM SIGOPSI, vol 36, p131-146, 2002.

[4] M. Mohania, S. Konomi, and Y. Kambayashi. "Incremental Maintenance of Materialized Views," In Proc. DEXA, pp. 551-560, 1997.

[5] A. Gupta and I. S. Mumick. "Maintenance of Materialized Views : Problems, Techniques, and Applications," In Proc. IEEE DEBU, Vol. 18, No. 2, pp. 3-18, 1995

[6] Timothy Griffin, Leonid Libkin, "Incremental Maintenance of Views with Duplicates", In Proc. ACM SIGMOD, p328-339, 1995

[7] Ashish Gupta, Inderpal Singh Mumick, V.S. Subrahmanian. "Maintaining Views Incrementally", ACM SIGMOD Record, Vol. 22, p157-166, 1993.

[8] Samuel Madden, Robert Szewczyk, Michael J. Franklin and David Culler. "supporting aggregate queries over ad-hoc wireless sensor network", Proc. IEEE WMCSA, p49-58, 2002.